February 13, 2025

Sent via ECF

Hon. Ona T. Wang Daniel Patrick Moynihan United States Courthouse 500 Pearl Street New York, NY 10007-1312

cc: All Counsel of Record Line (via ECF)

RE: The New York Times Co. v. Microsoft Corp., et al. No. 1:23-cv-11195 New York Daily News, LP, et al. v. Microsoft Corp., et al., No. 1:24-cv-03285 The Center for Investigative Reporting, Inc. v. OpenAI, Inc., No. 1:24-cv-04872

Dear Judge Wang:

Pursuant to the Court's January 24, 2025 Order (Dkt. 441), and on behalf of all parties in the above-referenced actions, counsel for OpenAI submits the below chart summarizing the status of various discovery disputes that were addressed at the January 22, 2025 conference.

Sincerely,

KEKER, VAN NEST & PETERS LLPⁱ

/s/ Andrew Dawson
Andrew Dawson
Counsel for OpenAI Defendants

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions				
	I. Disputes Related to Plaintiffs' Requests for Inspection										
305 ¹ 328	11/1/24 11/20/24	305 345	OpenAI's Training Data	News Plaintiffs	OpenAI	Resolved.	Joint Statement: The Parties have met and conferred, and no further action is required from the Court at this time. OpenAI ran Plaintiffs' "URL Search" of OpenAI training data and produced the search results to The Times and Daily News Plaintiffs on January 29, 2025. Plaintiffs are continuing to investigate these results. As directed by the Court, the parties met and conferred regarding proposed datasets and threshold values for the proposed "N-Gram Search" of OpenAI training data. OpenAI will attempt to run an agreed N-gram search on an identified set of training datasets and produce the results. If OpenAI encounters technical difficulties, it will raise them with Plaintiffs for further meet-and-confer discussions.				

¹ Unless otherwise indicated, all docket numbers are from *New York Times Company v. Microsoft Corp., et al.*, Case No. 1:23-cv-11195-SHS.

379	1/13/25	420	OpenAI's Log Output Data	News Plaintiffs	OpenAI	The Court heard argument at the January 22, 2025 conference, and ordered the Parties to meet and confer and inform the Court of the status of this issue in this joint chart. (Dkt. 441)	Plaintiffs' Position: OpenAI has refused to give The Times and Daily News Plaintiffs a clear answer as to what amount of output log data OpenAI deleted for each product at issue, and OpenAI will still not agree to preserve relevant output log data. Plaintiffs have reason to believe that OpenAI has deleted—and continues to delete—a substantial volume of relevant output log data based on OpenAI's representations to the Court about its privacy policies and deposition testimony from OpenAI's 30(b)(6) custodial witness. For example, with respect to API output data, OpenAI's general "default policy" is to delete data after 30 days. By contrast, for ChatGPT user conversations, OpenAI generally deletes data when requested by a user. Furthermore, although OpenAI's 30(b)(6) custodial witness testified that OpenAI has the ability to preserve output data that would ordinarily be slated for deletion under OpenAI's policies (which expressly permit OpenAI to retain data in order to comply with its "legal obligations"), OpenAI has only taken such steps with respect to specific accounts Defendants believe are referenced in the complaints. That is, OpenAI has only taken steps to preserve output log data for user accounts owned or controlled by Plaintiffs to conduct their pre-suit investigation—which OpenAI believes is helpful for its defenses—and is refusing to preserve output log data for third-party user accounts that support Plaintiffs' claims. The Parties met and conferred on February 7, 2025, to discuss potential technical means to preserve output log data relevant to the litigations, and OpenAI has not provided Plaintiffs with any specific technical proposal or the timing of any such proposal. Plaintiffs asked OpenAI to suspend its deletion policies on an interim basis while the Parties continue to meet and confer on possible technical solutions, and OpenAI has refused. In view of the relevance and importance of OpenAI's output log data to the issues in the case, Plaintiffs renew their request in ECF 379 for the Court to
-----	---------	-----	-----------------------------	-----------------	--------	--	---

reassert here—as they did in their motion—that OpenAI is "not
agree[ing] to preserve relevant output log data" and that "OpenAI is
refusing to preserve output log data for third-party user accounts that
support Plaintiffs' claims." That is not true. OpenAI's default policy is
to retain—that is, to preserve—all consumer ChatGPT conversations.
That is why OpenAI has retained billions upon billions of such user
conversations. See ECF 425 ¶ 4. Indeed, the amount of data makes
producing it all impractical, which is why the parties are in the midst of
exploring approaches to sampling the data. That OpenAI also respects
the privacy decisions of its users, including those that use the select
offerings that are subject to separate product-specific privacy policies
and contractual obligations (such as the API), does not change
the fact that, absent a user request to the contrary, OpenAI retains
consumer ChatGPT conversations.

The Court should decline Plaintiffs' improper request for the Court to reconsider its order at ECF 411. As explained below, and as required by the Court's order, the parties have been—and will continue to—engage in constructive conferrals to address the technological and privacy considerations implicated by Plaintiffs requests. That process is working and OpenAI looks forward to continuing to work with Plaintiffs constructively, rather than relitigating matters that the Court has already addressed.

At OpenAI's request, the parties met and conferred on February 7, 2025, as the Court directed, regarding the privacy and technological considerations implicated by Plaintiffs' request in the motion at ECF 379. Given the Court's directive at the last discovery hearing to focus on a path forward, OpenAI sought to discuss Plaintiffs' contention expressed in prior correspondence—that "OpenAI has tools and other mechanisms" that could address Plaintiffs' requests. OpenAI explained why, based on its investigation to date, the specific tool Plaintiffs identified did not function in a way that aligned with the request Plaintiffs were making. Plaintiffs asked a series of additional questions related to that tool and its capabilities, which OpenAI agreed to look into. OpenAI is currently investigating those highly technical questions, which is a necessary step in the process of determining what technical options might satisfy Plaintiffs' demands while at the same time taking account of the technological and privacy issues implicated by wholesale retention. At the same time, OpenAI is investigating the over two dozen other questions related to user conversation data that Plaintiffs have posed in a series of recent letters—matters the parties

		have been and will continue to address in correspondence and upcoming conferrals. In short, the conferral process the Court contemplated in its Order at ECF 411 is underway and should continue.

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
311	11/18/24	338	Microsoft's Log Output Data	News Plaintiffs	Microsoft	The Court heard argument at the January 22, 2025 conference, and ordered the Parties to meet and confer and inform the Court of the status of this issue in this joint chart. (Dkt. 441)	Plaintiffs' position: The Times and Daily News Plaintiffs sent a letter to Microsoft on February 3, 2025, outlining topics and questions to be addressed by Microsoft's experts during the informal technical discussion into discovery of Microsoft's log data, including questions about how Microsoft generated the sample log datasets, how Microsoft stores the log datasets, and the meanings of specific fields contained in Microsoft's log files. On February 10, 2025, technical experts for The Times and Daily News Plaintiffs attended an informal discussion with Microsoft's outside consultants. No Microsoft employees attended the discussion, and because Microsoft's outside consultants' knowledge of Microsoft's log data was limited to the same materials produced to Plaintiffs, the outside consultants could not answer many of the questions raised during the discussion and previously outlined in the February 3 letter. On February 12, The Times and Daily News Plaintiffs sent a letter to Microsoft to follow up on items that the consultants were not able to answer during the informal discussion, and requested responses by February 17. News Plaintiffs will continue to meet and confer with Microsoft on discovery into the log data. Microsoft's position: Microsoft's experts were prepared to answer the questions regarding the nature and accessibility of the data in Plaintiffs' February 3 letter, and, to the extent Plaintiffs asked those questions, Microsoft's experts provided responses. For instance, Microsoft's experts provided responses. For instance, Microsoft's experts provided responses, for instance, Microsoft's experts provided responses, to instance, Microsoft's experts had to aid in developing a sampling and querying methodology, including, but not limited to, describing what data is collected and what data is not collected, how the data is stored and can be access

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							(including Microsoft engineers) has complete knowledge of the data infrastructure and each of the hundreds of data fields and sub-fields as it is highly complex. To the extent Microsoft's experts were unable to answer each and every question during the meeting, that was largely the result of Plaintiffs' deviation from their provided list of questions and seeking information beyond the scope of the call. Microsoft has received Plaintiffs' February 12 letter, is investigating the "follow-up items" (most of which are new questions), and will provide responses to Plaintiffs in due course. Microsoft will continue to meet and confer with Plaintiffs on discovery into the log data.
				II. Disput	es Related to Cu	stodians and Cus	todial Documents
381	1/13/25	411	Microsoft's search terms (The News Plaintiffs request that Microsoft run additional search terms for its custodians' ESI, and that Microsoft extend the date range for collections to July 1, 2024)	News Plaintiffs	Microsoft	The Court heard argument at the January 22, 2025 conference, and ordered the Parties to meet and confer and inform the Court of the status of this dispute in this joint chart. (Dkt. 441)	Plaintiffs' position: The News Plaintiffs have made every effort to meetand-confer with Microsoft in good faith to resolve this dispute, including after the January 22, 2025 hearing. Consistent with this Court's instructions, the News Plaintiffs have narrowed their search term proposal to focus on key gaps in Microsoft's custodial review. But Microsoft's approach to search terms continues to violate the ESI Order, as well as this Court's recent directions. Microsoft should be ordered to apply the search terms in the News Plaintiffs' narrowed proposal made on February 6, 2025. Microsoft has violated the ESI Order and this Court's directions in at least three ways. First, between July and December 2024, the News Plaintiffs made six different search term proposals to Microsoft. Yet Microsoft accepted just 2 of the more than 100 search terms the News Plaintiffs requested, in violation of the ESI Order, which expressly contemplates a role for the receiving party in crafting search terms. Dkt. 251 at 1; see also Jan. 22, 2025 Hearing Tr. at 64:23-65:10 (counsel for Microsoft admitting that "we ran our own [search] terms" even though "The Times never agreed to our search terms"). Second, Microsoft took 25 business days (from December 30, 2024 until February 4, 2025) to return hit counts for the News Plaintiffs' December 30 proposal, which far exceeds the 3-business day turnaround required by the ESI Order. Dkt. 251 at 1. Third, in violation of both the ESI Order and this Court's instructions, Microsoft waited two weeks after the January 22 hearing to simply send the News Plaintiffs the list of search terms that Microsoft ran—and did so only after News Plaintiffs followed up five times. See Jan. 22, 2025 Tr. at 68:1-4 (instructing Microsoft to "put [the list of terms used]

Case 1:24-cv-03285-SHS-OTW Document 294 Filed 02/13/25 Page 8 of 14

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							together in one, send it to The Times, have the meet-and-confers in the next two weeks about whether there are gaps, and think about what you might want to do about those gaps"). Worse, Microsoft <i>still</i> refuses to provide hit counts for the search terms that it ran , in violation of the ESI Order. <i>See</i> Dkt. 251 at 1.
							The best path forward is to order Microsoft to apply the search terms in the News Plaintiffs' February 6, 2025 narrowed proposal, which reasonably yields approximately 370,000 additional documents for review. News Plaintiffs drafted this proposal immediately after receiving Microsoft's list of terms and designed it to address the substantial gaps in the search terms that Microsoft unilaterally selected. Indeed, Microsoft's terms were clearly crafted for the separate <i>Authors Guild</i> case, particularly because they include terms like "books1" and "books2", which are datasets comprised of book content. Meanwhile, Microsoft refuses to run terms designed to find documents about "generative search," a topic that Microsoft characterizes as unique to the News Cases. <i>See</i> Dkt. 148 at 2 (Microsoft arguing that "generative search" is a "separate new technology [that] underpins each of plaintiffs' counts in the Newspaper Cases" and that this technology is "not at issue" in the class actions). Not surprisingly, the News Plaintiffs have received almost no documents addressing the impact of generative search on news publishers. It is long past time to finish search term negotiations, and the News Plaintiffs' February 6 proposal fairly addresses the substantial gaps in Microsoft's custodial document review. Finally, News Plaintiffs disagree with Microsoft that resolving this dispute can be postponed until the next hearing in April. That approach undermines News Plaintiffs' goal to promptly begin depositions.
							Microsoft's position: As Microsoft has communicated to Plaintiffs, Microsoft is focused on working with Plaintiffs to move the issue of search term negotiations forward. To that end, following the January 30 conference, Microsoft provided to Plaintiffs (i) a comprehensive list of the search terms already used to collect documents from each of the 30
							agreed-on custodians; (ii) the hit report for Plaintiffs' previous search terms proposal, including adding in the extended date range Plaintiffs are requesting, which yielded approximately 650,000 unique parent-level items; and (iii) a hit report for its own, already-used terms for that

Case 1:24-cv-03285-SHS-OTW Document 294 Filed 02/13/25 Page 9 of 14

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							extended date range. In addition, Microsoft also provided Plaintiffs a search term hit report for Plaintiffs' February 6 proposal. The hit report revealed that the Plaintiffs' February 6 proposal yielded over 370,000 unique parent-level items. While this number is smaller than Plaintiffs' previous proposal, Microsoft estimates that the parent-level items in the most recent proposal will yield at least 600,000 (not 370,000) new documents to review once family documents are taken into account, and this is in addition to the more than 600,000—and counting—custodial documents that Microsoft has already reviewed. As of the filing of this update, Microsoft is preparing a line-by-line response, including counterproposals where appropriate, to Plaintiffs' February 6 list of 100 requested terms. In short, the parties are making progress in this area and should continue to meet and confer to finalize search terms.
							In the meantime, Microsoft disagrees with Plaintiffs' characterizations of its February 6 proposal, which does not "focus on key gaps." Plaintiffs complain that Microsoft's terms were crafted for the <i>Authors Guild</i> case and pick two terms, but Plaintiffs' selection of those two terms ignores the <u>dozens</u> of terms Microsoft used that target both issues common across cases (such as the training of OpenAI's models and the business relationship between Microsoft and OpenAI) and issues unique to the News cases that incorporate the products at issue, the News Plaintiffs themselves, and/or news. More importantly, Plaintiffs' proposal did not focus on "gaps," as the Court instructed. This will be the focus of many of Microsoft's line-by-line response to Plaintiffs' February 6 proposal.
							The February 6 proposal also includes many very general terms that make no attempt to target documents likely to be relevant to this case. As such, the best path forward is not to simply require review of the 600,000+ documents resulting from the search terms in the News Plaintiffs' February 6 proposal. Instead, and consistent with the Court's directive to meet and confer and fill the gaps, Microsoft is providing a line-by-line response, including counterproposals where appropriate, to Plaintiffs' February 6 list of 100 requested terms. Microsoft will provide its proposal by February 18, and Microsoft expects that following this the parties will be able to continue to narrow and resolve gap-filling disputes over the search terms in Plaintiffs' latest February 6

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							proposal, and the parties can raise any remaining disputes with the Court at the April 10 conference.
				III. Disp	oute Related to O	penAI's Documer	nt Productions
322- Sealed 323- Public	11/19/24	339	The Times's RFP 2 (documents OpenAI has submitted to domestic governmental entities relating to the allegations in the Complaint)	NYT	OpenAI	The Court heard argument on this motion at the January 22, 2025 conference, and ordered the parties to meet-and-confer and to provide a status update in this joint chart. (Dkt. 441)	Plaintiffs' position: The Times maintains its request for an order requiring OpenAI to identify any relevant domestic proceedings in which it has produced documents related to the issues in this case. But as a first step, OpenAI should produce any interrogatory responses or responses to written questions that OpenAI has submitted for the Delaware Attorney General's ongoing investigation into OpenAI's transition to a for-profit company. https://www.reuters.com/technology/artificial-intelligence/openais-altman-says-musk-offer-is-tactic-to-try-mess-with-us-2025-02-11/ (February 12, 2025 news article reporting that the Delaware Attorney General "is reviewing OpenAI's proposed changes to ensure the company is 'adhering to its specific charitable purposes for the benefit of the public beneficiaries, as opposed to the commercial or private interests of OpenAI's directors or partners.""). This approach targets highly relevant documents that have not otherwise been produced and are easily obtained. See A&M Recs., Inc. v. Napster, Inc., 239 F.3d 1004, 1015 (9th Cir. 2001) ("This 'purpose and character' element also requires the district court to determine whether the allegedly infringing use is commercial or noncommercial."). Yet OpenAI still refuses to even answer questions about this investigation, arguing that this information is irrelevant. Meanwhile, OpenAI appears intent on arguing that it was engaged in (and continues to be engaged in) a non-commercial use of The Times's work, making the commercial nature of OpenAI's business highly relevant. OpenAI below tries to sidestep the relevance of this issue by arguing the parties should be discussing The Times's other RFPs. The parties should

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							not get bogged down on whether the information requested here falls under RFP 2 alone or some other RFP as well. The Times is making a revised request for targeted information that is responsive to this RFP, and the requested information has not already been produced. Because OpenAI has been unwilling to reconsider its position, the parties remain at an impasse, and The Times respectfully asks the Court to adopt its proposed approach for this dispute, starting with this narrowed first step.
							OpenAI's position: The Court previously rejected The Times's motion to compel with regard to documents submitted to foreign regulators (ECF 398), and it should deny this parallel motion as well.
							The Times has issued 130 RFPs in this matter, which canvass all conceivably relevant topics. As relevant here, it has also requested all documents submitted to domestic governmental entities, without regard to whether those materials have been produced in response to other RFPs. The Court previously encouraged The Times to focus its discovery requests on "underlying information relevant to the case," not on materials that might have been deemed relevant in some separate context. (Transcript, Jan. 22, 2025, at 75:1-5). Indeed, with regard to alleged "commercial nature" of OpenAI's business, The Times has already issued an RFP focused on precisely that issue. <i>See</i> RFP 11. That RFP has been the subject of motion practice and extensive meetand-confer discussions. The Times should not be permitted to end-run around those negotiations simply by pointing to public reporting of a state regulator's entirely distinct inquiry. This is particularly true when the alleged relevance is a corporate change that has not even occurred yet, and where The Times blatantly misstates the alleged relevance - there is no dispute that OpenAI has had a for-profit entity since 2019. The Times is engaged in a fishing expedition, and its motion to compel should be denied.
							The Times's suggestion that it is offering a narrowed resolution to this RFP is false. It has clarified in meet-and-confer discussions that it expects OpenAI to canvas other, unidentified regulatory engagements for potentially relevant materials. Instead of casting about for materials potentially deemed relevant in other contexts, The Times should simply seek discovery of materials it believes to be relevant in <i>this</i> case.

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
DRI. NO.	Filed	No.	The Times's RFPs 61, 113-14 (documents concerning licensing negotiations and the market for			licrosoft's Docum	ent Productions Plaintiffs' position: This motion to compel addresses three categories of RFP disputes with Microsoft. The parties have continued to meetand-confer and have resolved category 3 (documents concerning the impact of Defendants' products on copyright owners) and a portion of category 1. M aterial disputes remain as to a portion of category 1 as well as category 2. For category 1, the parties have resolved RFPs 113-14, which seek documents concerning the market for the at-issue data. But disputes
<u>396</u>	1/13/25	<u>407</u>	model training data and the market for content for retrieval- augmented generation ("RAG") or	NYT	Microsoft	Partially Resolved: The Court heard argument on this motion at the January 22, 2025 conference, and ordered the Parties to meet and confer and to	remain as to RFP 61, which seeks documents concerning licensing negotiations. Specifically, The Times seeks documents concerning licensing negotiations for the at-issue generativeAI models and products that never resulted in a final license agreement. Consistent with the Court's instructions at the January 22, 2025 hearing, The Times made a narrowed proposal that focuses on licensing negotiations for news content. See Jan. 22, 2025 Tr. at 103:7-9 ("I think that you should narrow the scope of that meet-and-confer to what the market might be in the News cases."). The parties are continuing to meet-and-confer about this narrowed proposal, so The Times can understand the extent to which Microsoft possesses responsive documents and craft an appropriate demand.
			Microsoft's investments in OpenAI); The Times's RFP 109 (documents concerning the impact of Defendants' products on copyright owners).			provide an update in this joint chart. (Dkt. 441)	For category 2 (RFP 123), Microsoft continues to argue that it need not produce documents about its 2024 investments in OpenAI, claiming that 2024 conduct is somehow irrelevant since The Times filed the lawsuit in December 2023. That argument ignores the allegations of ongoing misconduct and Microsoft's ongoing liability for contributory and vicarious copyright infringement. FAC ¶ 66, 151, 169. Microsoft's 2024 investments in OpenAI are squarely relevant, and collecting these documents entails a minimal burden. Indeed, Microsoft has never opposed this discovery on burden grounds. To move things forward, Microsoft should produce documents about the \$750 million investment in OpenAI that it reportedly made in or around October 2024. https://www.cnbc.com/2024/10/30/microsoft-cfo-says-openai-investment-will-cut-into-profit-this-quarter.html . Microsoft should also

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							identify any additional 2024 investments in OpenAI. The Times will work with Microsoft to reach agreement on a reasonable scope for these productions, but Microsoft's blanket relevance objection has hindered these negotiations.
							Microsoft's position: Since the January 22 hearing, the parties have made progress in narrowing and resolving disputes over the requests for production at issue in Plaintiff's letter motion. Microsoft agrees with The Times that "category 3" (RFP No. 109) is resolved and "category 1" is partially resolved, specifically as to RFP Nos. 113 and 114 is resolved.
							With respect to the remaining portion of "category 1," RFP No. 61, it remains Microsoft's position that the Court's prior ruling on the same issue in <i>Authors Guild</i> should control, which the parties addressed in their letter briefs. Nonetheless, the parties should continue their discussion. The Times provided its position to Microsoft after the January hearing, Microsoft provided a response on February 7, and The Times responded on February 12; from that response, it appears the parties need to meet and confer further and possibly consult further with their clients.
							The parties have also had productive discussions with respect to with respect to RFP No. 123, or "category 2" as phrased by The Times. Microsoft raised legitimate questions regarding the relevance and scope of what The Times seeks and is considering The Times's most recent position, which was provided on February 12. Microsoft's written objections contained precisely these objections, and in the meantime, Microsoft has sought to understand The Times's position as to relevance because, the relevance informs what discovery is proportionate to the needs of the case. Microsoft does not agree with The Times's assertion that collecting the requested documents entails minimal burden, despite the fact that Microsoft has never opposed the discovery on burden grounds. The Times's position suggests that Microsoft is somehow precluded from raising issues as to both relevance and burden, but these issues are appropriate for addressing in a good faith meet and confer process. In any event, the parties are continuing to meet and confer regarding this request and Microsoft

Motion Dkt. No.	Date Filed	Opp. Dkt. No.	Topic in Dispute	Party Raising Dispute	Party Responding	Status of Dispute	Parties' Shorthand Positions
							expects that the parties will be able to reach an agreement as to at least some portion of this request, if not all of it.

ⁱ All parties whose electronic signatures are included herein have consented to the filing of this document.